

Open algorithms, social justice, and predictive policing

Daniel Krashen, Professor of Mathematics, Rutgers University

We live in a world in which algorithms are broadly used by companies and organizations which play important roles in our society.

When these algorithms are secret, for example, when they are protected by intellectual property laws, or when their use is not disclosed, we are unable to scrutinize their validity, to check them for unintended (or intended) biases, or to ensure their correct application. On the other hand, when algorithms and their use are made publicly available, and exist within open scientific, mathematical or statistical dialogue, we have a fighting chance of engaging with them, and making them better as a community.

Cathy O'Neil has led the effort to raise awareness of bias in real world applications of mathematics. She has helped clarify that we, as mathematicians and scientists, have a collective responsibility to try and ensure that algorithms and methods are developed and used in a way which promotes transparency so that we can locate and try to correct these biases.

Predictive policing is an important case to consider. Broadly speaking, this is the process of constructing algorithms to predict where crimes are likely to occur, and has been used to direct patrolling behavior of police officers with the intention of reducing criminal behavior. While I can't speak for all of the companies and algorithms, some of these algorithms in use are publicly available in the scientific literature, have been the object of a great deal of study, and many people have scrutinized them and worked towards their improvement to address potential problems of inequity and bias.

Police patrolling will not simply end. If mathematicians, scientists and others don't come together to help formulate algorithms about patrolling, we can do little to influence the potential bias that the police can (and likely will) bring. But if the algorithms used by the police are transparent, and placed in a forum of public scientific discussion, we can work together to find potential sources of bias and

inequity, and address them. If the algorithms aren't available, if the police obtain them through businesses that keep them confidential, this conversation can never happen, and this is when society will really suffer.

I don't think this is the time for academics to walk away from the conversation with the police or with other institutions and companies, but rather now is the time to go deeper, to analyze how particular algorithms are used, to push for maximal transparency, and try and identify and correct bias where we can. This is the time to engage with our colleagues, who have developed and refined their expertise to think deeply about these problems, and who have developed a dialog with various social institutions. They are an incredibly valuable resource which should not be squandered. If we are concerned about their results, assumptions, or the application of their work, we need to engage with them first in the realm of scientific discourse about their research, of which a great deal is publicly available.

I should make it clear that this is not my own area of expertise (which is noncommutative algebra and arithmetic geometry). Nonetheless, I have spent some time recently trying to familiarize myself with the literature, to the extent that I have been able. I'd like to point out some sources I found particularly informative on this topic:

Cathy O'Neil's book, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy,

- https://www.amazon.com/dp/0553418831/ref=cm_sw_em_r_mt_dp_U_OnL7EbPB39V9B

Here's a nice article on some of the theory and practice of predictive policing algorithms as used by PredPod:

- <http://dx.doi.org/10.1080/01621459.2015.1077710>
(https://www.researchgate.net/publication/282772661_Randomized_Controlled_Field_Trials_of_Predictive_Policing)

And this is a useful summary and overview of predictive policing as a whole:

- <https://nij.ojp.gov/library/publications/predictive-policing-role-crime->

forecasting-law-enforcement-operations

This article points out potential for biases being reinforced by the algorithm in "feedback loops":

- <https://doi.org/10.1111/j.1740-9713.2016.00960.x>

While important concerns were brought up, this article seems to have started a good deal of public discourse about potential negative effects of predictive policing, most of which seemed unrelated to the actual algorithm in question or its use. I found the following summarizing article very interesting:

- <https://theappeal.org/the-truth-about-predictive-policing-and-race-b87cf7c070b1/>

Some articles examining whether or not these potential biases of PredPol's algorithm seem to arise in practice:

- <https://www.tandfonline.com/doi/full/10.1080/2330443X.2018.1438940>
- <https://kb.osu.edu/handle/1811/85819>

An article addressing a strategy to fix the feedback effects and remove the problem of reinforced bias:

- <http://proceedings.mlr.press/v81/ensign18a.html>

Many more articles of interest and further details can be found by simply going to the websites of the various authors of these papers, and looking at their publications.

We have been fortunate, as a community, that some of these algorithms are available in the scientific literature, and can be further explored. I hope that we can move to make things more open in the future, so that more algorithms which affect our everyday lives can be made transparent and available, and their uses may be fully disclosed. This is something that we need to work towards as a society. And as scientists, we need to engage with the algorithms that do exist, to test them, to critique them, and work to fix them when needed.

